PLANNING MALAYSIA:
Journal of the Malaysian Institute of Planners
VOLUME 19 ISSUE 2 (2021), Page 72 – 84

APPLICATION OF MACHINE LEARNING IN ANALYSING HISTORICAL AND NON-HISTORICAL CHARACTERISTICS OF HERITAGE PRE-WAR SHOPHOUSES

Nur Shahirah Ja'afar[1] & Junainah Mohamad[2]

[1]*Faculty of Architecture, Planning and Surveying*
UNIVERSITI TEKNOLOGI MARA, SHAH ALAM, MALAYSIA
[2]*Department of Built Environment Studies & Technology, Faculty of Architecture, Planning and Surveying*
UNIVERSITI TEKNOLOGI MARA, PERAK BRANCH, MALAYSIA

Abstract

Real estate is complex and its value is influenced by many characteristics. However, the current practice in Malaysia shows that historical characteristics have not been given primary consideration in determining the value of heritage properties. Thus, the accuracy of the values produced is questionable. This paper aims to determine whether the historical characteristics of the pre-war shophouses at North-East Penang Island, Malaysia contribute any significance to their value. Several Machine Learning algorithms have been developed for this purpose namely Random Forest, Decision Tree, Lasso Regression, Ridge Regression and Linear Regression. The result shows that the Random Forest Regressor with historical characteristics is the best fitting model with higher values of R-squared (R^2) and lowest value of Root Mean Square Error (RMSE). This indicates that the historical characteristics of the heritage property under study contribute to its significant value. By considering the historical characteristics, the property's value can be better predicted.

***Keywords**:* Pre-war shophouses, machine learning, historical characteristics, random forest, price prediction

[1] Postgraduate student at Universiti Teknologi MARA. Email: sirashahirah@gmail.com

INTRODUCTION

In 2008, George Town Penang Island Malaysia was recognized as a historic city inscribed by the UNESCO World Heritage Site (Azizan & Desa, 2020). This recognition was granted due to the prevalence of diverse tangible and intangible heritage surrounding the George Town area (Rahman, 2018). This paper focuses on the cultural heritage site which is made up of pre-war shophouses located within the core and buffer zones in George Town. Being labelled as a tangible cultural heritage, the property's value is of historical significance. Hence, there is a crucial need to conserve and preserve this heritage property as part of an economic indicator. According to Ruijgrok (2006), the preservation of cultural heritage produces considerable benefits in terms of economic and financial value. The total economic value of cultural heritage sites consists of their use-value and non-use value (Palanca-Tan, 2020). In terms of economic value, the heritage pre-war shophouses have a use value due to their status as private property. The owners are the direct beneficiary of the value due to their direct consumption of the property (SGS Economics and Planning, 2017).

In real estate valuation, it is important to estimate the possible value of a property in order to achieve the accurate value (Mohamad & Ismail, 2019). In the study by Mohamad, Ja'afar, and Ismail, (2020), the accurate valuation of a heritage property is indicated to facilitate better decision making. Thus, others would appreciate, acknowledge and assist in the management of heritage property so as to produce a proper and reliable method for their valuation.

Mohamad, Ismail and Rahman (2015) questioned about the proper historical characteristics that should be considered when valuing heritage property. Based on the study by Mohamad and Ismail, (2017); Zin and Ismail, (2019), there are several characteristics influencing the value of heritage property such as property transaction characteristics (e.g. tenure, lot number, building number, share, year of valuation), structural characteristics (e.g. main floor area, building improvement, types of floor, wall material, maintenance inside and outside) and historical characteristics (e.g. façade status, ensemble and authenticity). Ismail and Zin, (2019) recommended a proper and updated of historical characteristics in valuing heritage property so as to enhance the accuracy of existing heritage property valuation methods. The study by Ja'afar & Mohamad (2020) took into consideration historical characteristics by using the multiple regression analysis with higher R^2. The aforementioned studies revealed the importance of considering historical characteristics in valuing pre-war shophouses.

In the existing studies, several approaches had been used for estimating real estate property such as the sales comparison method, contingent valuation method, hedonic pricing method and travel cost method; while in the industry, the approaches used include the sales comparison method, income method and cost method (Ja'afar & Mohamad, 2020; Ruijgrok, 2006). This study aims to estimate the tangible heritage property value of pre-war shophouses; thus, the researchers need to identify the relevant historical characteristics for estimating the value price of the heritage property as there are no solid evidence of a proper valuation method for heritage property (Ja'afar & Mohamad, 2020). Besides that, authors also suggest observing the collected heritage characteristics using the most current and widely used statistical technique namely machine learning. Authors also pointed out that machine learning had been mostly used in estimating the value of residential properties and yet no study has been conducted in estimating the value of privately-owned heritage properties specifically in Malaysia, thus the study of heritage property using machine learning remains undiscovered. To study the heritage property's characteristics, the researchers had to identify which historical characteristics could give the most significant value to the pre-war shophouses using machine learning.

According to Baldominos and Blanco (2018), through the use of machine learning, the hidden value in data sources can be analysed to derive actionable insights from the data. The application of machine learning can also help identify more opportunities in the real estate market. This would drive real estate investors to give more attention to a property's surroundings. Thus, this study attempts to identify the physical aspects of heritage property of which characteristics are significant and should be considered in heritage property valuation, besides focusing on the historical group features which could facilitate competent property practitioners in determining the ways in which to determine heritage property values.

RESEARCH BACKGROUND

Heritage

According to the National Heritage Act 645 (2005) (NHA), the generic meaning of "heritage" entails sites, objects and underwater cultural heritage. This current paper focuses on pre-war shophouses that are labelled as cultural heritage. There are several definitions of cultural heritage. Gabriel (2020) defined it as the handiwork of humans that are deemed worthy of preservation and of which can be tangible or intangible. Thus, pre-war shophouses are known as tangible cultural heritage following the decree of the NHA which states that an area, monument or building in this category is deemed as tangible cultural heritage. UNESCO recognized the pre-war shophouses located within George Town as a World Heritage Site because the buildings possess all the criteria mentioned in the Outstanding Universal Value (OUV). The criteria include: (i) representing the multi-cultural trading town which involve Malay, Chinese, Indian and European cultures with different architectures, technologies and monument arts, (ii) representing multi-cultural traditional living influences be it tangible or intangible aside from the existence of various religious buildings, languages, foods, daily life habits and ethnicities, and (iii) reflecting the various cultural architectures from the Malay Archipelago, India, China and Europe via the presence of unique building types and cultures (Azizan & Desa, 2020; Foo & Krishnapillai, 2019). Figure 1 shows the classification of heritage sites in Malaysia. This study identifies significant historical characteristics through empirical and theoretical study using machine learning modelling.

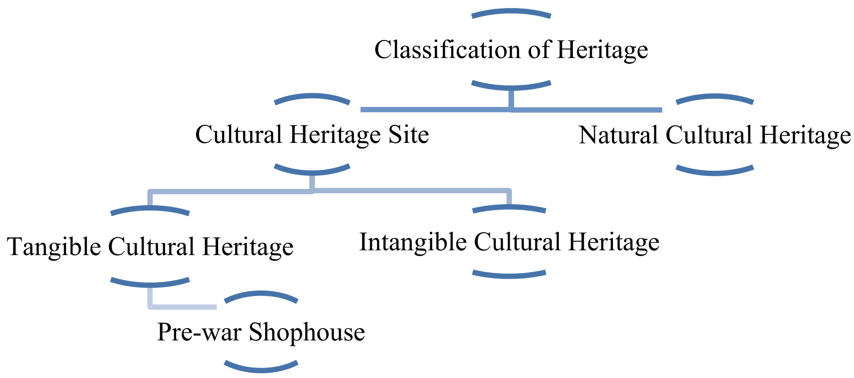

Figure 1: Classification of Heritage Site

Factors Affecting Pre-War Shophouses
Historical characteristics
Every heritage building has its own valuable historical characteristics which in turn represent an image of its prestige and attraction to the property (Azizan & Desa, 2020). There are various ancient cultures throughout human history; hence, researchers cannot simply ignore historical characteristics in valuing heritage property. This study examines the historical characteristics of the George Town World Heritage Site Incorporated (2016) pre-war shophouses in George Town which has six heritage shophouses from different periods with different facades and characteristics. In the context of this current study, the historical characteristics taken into consideration include transacted price, year of transaction, street of property, storey, land area, main floor area, roof material, floor material, ceiling material, maintenance inside and outside, historical styles, five-footway/walkaway, multifunction building, architectural, functionalistic, ensemble and authenticity (Mohamad & Ismail, 2019; Mohamad et al., 2020; Yeow Wooi, 2015). These historical characteristics are used in the modelling using the machine learning statistical technique.

Non-historical characteristics
Non-historical characteristics refer to characteristics with no historical value or heritage context. Normally, in valuing real estate properties like housing or private shophouses, the characteristics that will be taken into consideration include age of building, location of property, land area, storey, tenure, neighbourhood and main floor area (Zulkarnain, 2020). These characteristics are valued based on the conditions and sales evidence from the open market on the property types; hence, they are different from historical characteristics which cannot be easily compared to other properties (Shipley, 2000).

Machine Learning
Machine learning has become a popular programming practice among researchers for solving problems by predicting current existing data gathered from past data records (Milutinovic, 2019). There are two categories of problem-solving using machine learning namely supervised learning and unsupervised learning algorithms (Fiorucci & James, 2020). Between the two, the most commonly used is the supervised learning algorithm specifically for predicting ϒ (Horino & Nonaka, 2017). The supervised learning algorithm is used for predicting the outcome of a given input; it uses the examples of the input or output pairs and requires human effort to create a training set for building the machine learning algorithm process (Jordan & Mitchell, 2015). Unsupervised learning algorithms are different because they have no known output and no instructor to instruct the learning algorithms; they also involve the analysis of unlabelled data from assumption i.e., the extraction of information from the input data to build algorithms (Muller & Guido, 2017).

The machine learning approach has been used in the real estate field for several years. It is used to determine the market value of buildings, predict long term values, match profiles, and generate real estate listings and other information related to property forecasting (Phan, 2019). However, the machine learning approach is seldomly employed in the real estate industry in Malaysia. Therefore, in order to assess any benefits that could be gained from it, researchers apply it as a price prediction. There are five most commonly used algorithms in real estate analysis namely Random Forest, Decision Tree, Linear Regression, Lasso Regression and Ridge Regression as shown in Figure 2.

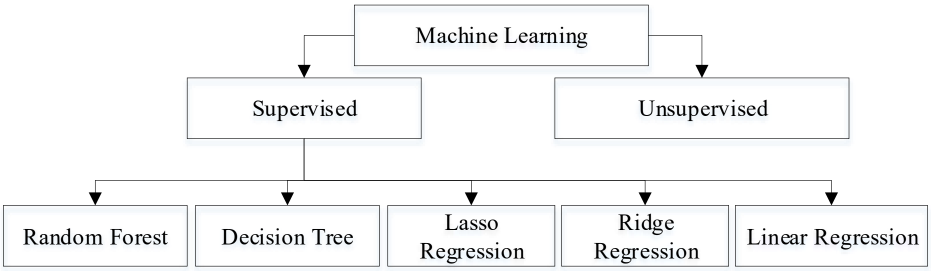

Figure 2: Machine Learning Algorithms Diagram

According to Mohamad et al. (2020), the most successful supervised machine learning algorithm as identified in various studies is the Random Forest algorithm. Random Forest is also known as an ensemble learning which can be used in classification and regression problem methods; one of its advantages is protecting against overfitting which in turn improves performance (Sabbeh, 2018). It is usually used as a decision-making tool in real estate specifically for predicting the price of housing (Horino & Nonaka, 2017). Random Forest has a Decision Tree collection known as "Forest", but Random Forest generalizes better than Decision Tree towards improving accuracy i.e., by selecting the highest votes (Fiorucci & James, 2020).

Next is the Linear Regression algorithm or better known as ordinary least square OLS (Varma & Sarma, 2018). Linear Regression is used to estimate the cost of houses, total sales and number of calls. There are two types of Linear Regression namely simple linear regression and multiple linear regression. Simple linear regression uses one independent variable while multiple linear regression employs more than one independent variable. Basically, Linear Regression is applied for predicting, forecasting and studying the relationship between two variables. Linear Regression equation is $Y = a + \beta x$ where Y is the market price of property and x is the given input (Borde & Rane, 2017).

Next, the Decision Tree algorithm is used in classification and regression problems but mostly for classification. It is used for visualizing the decision-making process. More commonly, it is used for selecting variables, accessing the significant connections between variables, monitoring missing values, and predicting data and management (Song & Lu, 2015). According to Sabbeh (2018), Decision Tree generates a tree-like structure which comprises internal nodes, branches and leaf nodes and represents the decision model.

Next is the Lasso Regression algorithm, which stands for Least Absolute Shrinkage and Selection Operator (Satish & Rao, 2019). According to Shinde and Gawande (2018), Lasso Regression is known as an L1 regularization technique because it is one of the most powerful formulas in regression, it works by reducing the error between the predicted and actual observations. Last but not least is Ridge regression, a linear model for regression which uses the same formula used in predictions using OLS (Muller & Guido, 2017). This algorithm is able to fit an additional constraint called regularization during training data. Ridge regression is known as L2 regularization which prevents overfitting during training. In modelling, the evaluation formula for determining the performance metrics used root mean square error and R-square to show the good values predicted by the algorithm performance (Mohamad et al., 2020).

Machine learning has numerous benefits. It is classified as a continuous improvement approach because most machine learning algorithms are capable of learning from available data and constantly provides new knowledge as well as enhances the accuracy and efficacy of decision making with subsequent training. In addition, it helps in recognizing patterns and trends of data in huge volumes and in discovering patterns and trends that are not obvious to humans i.e., by browsing and purchasing history of data to disclose the finding (Levantesi & Piscopo, 2020). However, there are also flaws in machine learning such as the necessity to acquire large volumes of high quality and accessible data to train on. Occasionally, there is a need to wait for the production of new information as the machine learning algorithms require adequate time to learn and improve with a significant amount of precision and relevance to fulfil their purpose (Borde & Rane, 2017; Raschka, 2020; Voutas Chatzidis, 2019).

METHODOLOGY

Dataset

To compensate for the lack of literature on historical characteristics and other related criteria, researchers have collected several historical characteristics from empirical and theoretical study regarding the pre-war shophouses. There 19 heritage pre-war shophouses characteristics used in the machine learning modelling such as transacted price, year of transaction, street, storey, land area, main floor area, roof material, floor material, wall material, ceiling material, maintenance outside, maintenance inside, multifunction, five-footway, architectural functionalistic, historical styles, ensemble, authenticity and position. Dataset for empirical observation was collected from the National Property Information Centre and inspection, while the theoretical study was taken from a book entitled "George Town Historic Cities of the Straits of Malacca Special Area Plan" (George Town World Heritage Incorporated, 2016; Gwynn Jenkins, 2013). These books were supported by the Penang Island City Council, Penang Town and Rural Planning Department and World Heritage Organization (WHO). As the property is a tangible cultural heritage, the researchers had to differentiate between the historical characteristics and non-historical characteristics in modelling using machine learning so as to determine the variables that could affect the heritage property's price.

Selected Variables

To train the variables using machine learning, the researchers divided the variables into three groups called features. The first group is called "All Features" comprising all the historical characteristics and non-historical characteristics of the properties to estimate the extent of their significance. The second group is named "Historical Features" comprising only the historical characteristics of the properties. The third group is called "Non-Historical Features" comprising only the non-historical characteristics of the properties. Through these different features used in training, the researchers can observe the outcomes of the machine learning at the end of the process.

Table 1: Features in Machine Learning

All Features (19 variables)	Transacted Price, Year of Transaction, Street, Storey, Land Area, Main Floor Area, Roof Material, Floor Material, Wall Material, Ceiling Material, Maintenance Outside, Maintenance Inside, Multifunction, Five-Footway, Architectural, Functionalistic, Historical Styles, Ensemble, Authenticity, Position.
Historical Features (18 variables)	Transacted Price, Year of Transaction, Street, Storey, Land Area, Main Floor Area, Roof Material, Floor Material, Wall Material, Ceiling Material, Maintenance Outside, Maintenance Inside, Multifunction, Five-Footway, Architectural, Functionalistic, Historical Styles, Ensemble, Authenticity.
Non-Historical Features (8 variables)	Transacted Price, Year of Transaction, Street, Storey, Land Area, Main Floor Area, Position.

Models Configuration

As mentioned earlier, this paper will predict an outcome from the given variables. Here, the researchers will identify the features that could influence the property's price by using the performance metrics. Thus, the machine learning model was built from the given features into training sets before the evaluation of the models. This model entails several processes refer Figure 3.

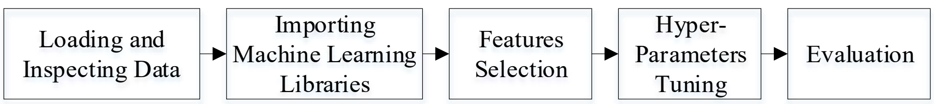

Figure 3: The Configuration in Modelling the Machine Learning Algorithms

According to Figure 3 above, there are five steps in modelling the machine learning in this paper. The steps are as follows:

1. Upload the files containing dataset (csv files) for inspecting the data in Python platform to loading and inspecting data.
2. Import machine learning libraries by calling Sklearn or Scikit-Learn libraries.
3. Select the features as shown in Table 1. The researchers had divided the variables into three groups namely: (1) All features, (2) Historical features, and (3) non-Historical features. This is to help determine the variables that influence the property's price.
4. Conduct the Auto Hyper-Parameter Tuning, by calling the *"best_estimator"* in Scikit-Learn library to help optimize the tuning configuration.
5. Lastly, conduct the evaluation process by analyzing the performance metrics using the R^2 and Root Mean Square Error. The result will show the selected algorithm by referring to the best performance metrics in Table 2.

RESULTS

After modelling the three features, the machine learning produced the correlations between the features and the results of the features with the selected algorithms. Each feature was generated with the chosen algorithms which were evaluated using the performance metrics i.e., the R^2 and the RMSE. The Root Mean Square Error is a standard way for measuring how errorless the model is in predicting quantitative data; meanwhile, the R^2 indicates the significant effects on the dependent variable (Ho, 2020; Yilmazer, 2020).

The correlation results of each properties characteristic calculated by machine learning are land area = 0.60, year of transaction = 0.38, main floor area = 0.35, position = 0.30, architectural functionalistic = 0.23, storey = 0.18, wall material = 0.13, historical styles = 0.12, street = 0.08, roof material = 0.04, ensemble = 0.018, five-footway = 0.009, multifunction = 0.0077, authenticity -0.061, ceiling material = -0.0814, maintenance inside = -0.13, floor material = -0.14 and maintenance outside = -0.15. The results show that there are correlations between the dependent and independent variables of property characteristics. The dependent variable is transacted price, and the remaining eighteen are the independent variables. The highest independent variable is land area while the lowest independent variable is maintenance outside. From this result, we can see that each characteristic has its own connection in influencing the property's price prediction.

Table 2: The Result of Features using Machine Learning

No	Algorithm	All Features		Historical Features		Non-Historical Features	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
1.	Random Forest	0.917	415258.7	0.971	242625.4	0.912	427367.7
2.	Decision Tree	0.821	611493.4	0.833	590643.1	0.899	459491.8

3.	Lasso	0.807	635270.7	0.822	610198.8	0.799	648317.9
4.	Ridge	0.804	640737.7	0.604	910646.3	0.799	637353.4
5.	Linear	0.380	746071.9	0.328	776845.2	0.806	648316.1

As presented in Table 2, there are three different features namely "All Features", "Historical Features" and "Non-Historical Features" together with different results. For the "All Features", the selected algorithm was Random Forest based on the performance metrics with the highest R^2 which is 0.917 and the lowest Root Mean Square Error which is 415258.7. This feature comprises all the property characteristics as shown in Table 1. Next, "Historical Feature" also nominated Random Forest as the most suitable for this dataset based on the performance metrics with the highest R^2 at 0.971 and the lowest Root Mean Square Error at 242625.4. Lastly for "Non-Historical Features", once again Random Forest was shown to be the most suitable algorithm based on the performance metrics of R^2 and Root Mean Square Error of 0.912 and 427367.7.

DISCUSSION

From the finding, the best algorithm is Random Forest model, because the result of R^2 and Root Mean Square Error are the best within other algorithms. The significant contributions by Random Forest model are through the result of accuracy. Accuracy is an indication of prediction price values, if prediction price were correctly formulated and modelled to achieve a fairly high level of accuracy, the level of prediction values will be the higher one. Thus, the accuracy achieved by Random Forest model is the higher one so it was formulated and modelled correctly.

In addition, data of features also contribute to the high accuracy. Among the results of features, "Historical Features" were selected as the best features dataset due to its accuracy from the Random Forest model. Dataset of "Historical Features" contain more variables than "Non-Historical Features". The same variables in both features are transacted price, year of transaction, street, storey, land area and main floor area only and the rest features are different. Thus, the differences of additional variables influence the accuracy result of models.

In conclusion, the "Historical Features" performance metrics is better than the other features hence indicating that historical characteristics can influence the performance of price prediction. For future studies, historical characteristics shall be considered in price prediction apart from upgrading the crucial characteristics that visualize the image of Historical City.

CONCLUSION

This study applied five algorithms namely Random Forest, Linear Regression, Decision Tree, Lasso Regression and Ridge Regression on the heritage property dataset. Within these algorithms, the findings showed that the model which best fits the data condition in modelling was Random Forest based on the performance metrics of the R^2 and the Root Mean Square Error. The best result for Random Forest was from the historical feature or historical characteristics data based on the highest value of R^2 and lowest value of Root Mean Square Error. Through empirical study, it helps in reviewing and restoring the latest recognition of historical characteristics areas for improvement of their special significance. By studying the different groups of property characteristics, the researchers had addressed the gap by presenting the empirical results from different property characteristics using the machine learning technique. However, machine learning also has its limitations including its need for high volumes of datasets for learning and predicting (Raschka, 2020). Another limitation is the lack of verifications on the standard of historical characteristics that should be considered in valuation. Future researches on historical characteristics are encouraged to improve or upgrade the historical characteristics based on the property's situation. This study is beneficial for the property owners in terms of income generation from the sale and purchase of their heritage properties. For the public, this study facilitates in defining its cultural identity apart from contributing to the preservation and conservation of historical characteristics (Armitage & Irons, 2013).

ACKNOWLEDGEMENT

The researchers would like to express their gratitude to the National Property Information Centre (NAPIC) for providing the property heritage data in Penang Island, and also the Malaysian Ministry of Education (MOE) for funding this research (FRG: FRGS/1/2018/WAB03/UITM/03/1).

REFERENCES

Armitage, L., & Irons, J. (2013). The Values of Built Heritage. *Journal of Property Management*, *31*(3), 246–259.

Azizan, M. A., & Desa, H. (2020). The Challenges in Conservation Practices in Malaysia: UNESCO Heritage Site, Georgetown. *Journal of Science*, *2213*, 1–2.

Baldominos, A., & Blanco, I. (2018). Identifying Real Estate Opportunities Using Machine Learning. *Journal of Applied Science*, *8*(2321), 1–23.

Borde, S., & Rane, A. (2017). Real Estate Investment Advising Using Machine Learning. *Research Journal of Engineering and Technology*, *04*(03), 1821–1825.

C. Muller, A., & Guido, S. (2017). *Introduction to Machine Learning with Python: A Guide for Data Scientists* (First Edit; D. Schanafelt, Ed.). New York City.

Fiorucci, M., & James, S. (2020). Machine Learning for Cultural Heritage: A Survey. *Journal of Pattern Recognition Letters*, *133*, 102–108.

Foo, R., & Krishnapillai, G. (2019). Preserving the intangible living heritage in the George Town World Heritage Site. *Journal of Heritage Tourism*, *14*(4), 358–370.

Gabriel, S. P. (2020). *Making Heritage in Malaysia* (First Edit; S. P. Gabriel, Ed.).

Singapore: Springer Nature Singapore Pte Ltd.
George Town World Heritage Incoperated. (2016). *Special area plan*.
Gwynn Jenkins, T. Y. W. (2013). Heritage Building Materials. *George Town.*
Ho, W. K. O. (2020). Predicting property prices with machine learning algorithms. *Journal of Property Research*, 1–23.
Horino, H., & Nonaka, H. (2017). Development of an Entropy-Based Feature Selection Method and Analysis of Online Reviews. *Journal of Engineering*, 2351–2355.
Ismail, S., & Zin, N. M. (2019). A Systematic Review on Spatial-Based Valuation Approach for. *International Journal of Real Estate Studies*, *13*(2), 59–70.
Ja'afar, N. S., & Mohamad, J. (2020). An Assessment of Heritage Property Values using MRA: George Town. *Journal of Sustainable Environmen*, *7*(2), 37–60.
Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, Perspectives and Prospects. *Journal of Electrical Engineering and Computer*, *349*(6245), 255–260.
Levantesi, S., & Piscopo, G. (2020). The importance of economic variables on London real estate market: A random forest approach. *Risks*, *8*(4), 1–17.
Milutinovic, M. (2019). Towards Automatic Machine Learning Pipeline Design. University of California, Berkeley.
Mohamad, J., & Ismail, S. (2017). Assessment of Heritage Property Values Using MRA and RTR. *Journal of Environment Behaviour*, 207–219.
Mohamad, J., & Ismail, S. (2019). Capabilities of Revealed Preference Method for Heritage Property Valuation. *Journal of the Malaysia Institute of Planners*, *17*(1), 377–379.
Mohamad, J., Ismail, S., & Rahman, R. A. B. (2015). Valuers' Perception on The Current Practice of Heritage Property Valuation. *Real Estate Society*, 18–21.
Mohamad, J., Ja'afar, S., & Ismail, S. (2020). Heritage Property Valuation using Machine Learning Algorithms. *Annual Pacific Rim Real Estate Society*, 1–12.
Palanca-Tan, R. (2020). Willingness to Pay Urban Households for the Conservation of Natural Resources in Rural Area use CVM. *Journal of Science*, *149*(2), 363–373.
Phan, T. D. (2019). Housing Price Prediction using Machine Learning Algorithms: The Case of Melbourne City. *International Conference on Machine Learning*, 35–42.
Rahman, S. (2018). Emerging Built Heritage Commodification of Boutique Hotels in WHS Penang. *Journal of the Malaysian Institute of Planners*, *16*(4), 104–116.
Raschka, S. (2020). Machine learning in python: Main developments and technology trends in data science of machine learning. *Information (Switzerland)*, *11*(4).
Ruijgrok, E. C. M. (2006). The Three Economic Values of Cultural Heritage: A Case Study in the Netherlands. *Journal of Cultural Heritage*, *7*, 206–213.
Sabbeh, S. F. (2018). Machine-learning techniques for customer retention: A comparative study. *Journal of Advanced Computer Science*, *9*(2), 273–281.
Satish, G. N., & Rao, M. D. S. (2019). House Price Prediction Using Machine Learning. *Journal of Innovative Technology and Exploring Engineering*, *8*(9), 717–722.
SGS Economics and Planning. (2017). *The Value of Heritage: Summary Report*.
Shinde, N., & Gawande, K. (2018). Valuation of House Prices using Predictive Techniques. *Journal of Advances in Electronics Computer Science*, *5*(6), 34–40.
Shipley, R. (2000). Heritage Designation and Property Values: Is there an effect? *International Journal of Heritage Studies*, *6*(1), 83–100.

Song, Y., & Lu, Y. (2015). Decision Tree Methods: Applications for Classification and Prediction. *Journal of Biostatistics in Psychiatry*, *27*(2), 130–136.

Varma, A., & Sarma, A. (2018). House Price Prediction Using Machine Learning and Neural Networks. *Journal of Computational Technologies*, 1936–1939.

Voutas Chatzidis, I. (2019). *Prediction of Housing Prices based on Spatial & Social Parameters using Regression & Deep Learnin*. University of Thessaloniki.

Yeow Wooi, T. (2015). *Penang Shophouses a Handbook of Features and Materials* (First Edit; R. Gareth & G. Jenkins, Eds.). George Town.

Yilmazer, S. (2020). A Mass Appraisal Assessment Study using Machine Learning Based on Multiple Regression and Random Forest. *Land Use Policy*, *99.* 104889.

Zin, N. M., & Ismail, S. (2019). Critical Determinants of Heritage Property Value: A Conceptual Framework. *Journal of the Malaysian Institute of Planners*, *17*(1), 219–231.

Zulkarnain, S. H. (2020). Structural, Locational and Environmental Attributes effects the Residential Property Value in Flood Risk Area. IOP Conference Series: *Earth and Environmental Science*, *497*. 012017

Received: 17th May 2021. Accepted: 9th July 2021

